



## Privacy-preserving AI in Systems Medicine with Federated Learning

2nd ECSCI Workshop April 27, 2022

#### **Julian Matschinske**

M.Sc. Computer Science, Lead Developer / PhD Student

Prof. Dr. Jan Baumbach

Institute for Computational Systems Biology

University of Hamburg



#### **FeatureCloud Project**



- EU-funded Horizon 2020 project
- 8 partners from Austria, Denmark, Germany, the Netherlands and Romania
- Started January 2019







### **Central Machine Learning (ML)**



FeatureCloud concept: Hospital's cloud communication present situation

- Each datapoint sent to central instance
- Model computed completely centrally





#### **Federated ML**



FeatureCloud concept: Hospital's cloud communication with FeatureCloud

- Intermediate models are trained locally
- Final model aggregated centrally





#### **Unified Platform for Collaborative Al**







## **Genome-wide Association Studies (GWAS)**

- Case/control studies on SNP data
- Each SNP examined individually (LR)
- Find SNPs related to diseases







#### **GWAS Cont'd**







#### **sPLINK: Federated Tool for GWAS**



#### - URL: exbio.wzw.tum.de/splink/

- Reza Nasirigerdeh et al. (published in Genome Biology)





#### **sPLINK vs. Aggregated and Meta-Analysis**







#### **sPLINK: Accuracy**

#### Obtains identical results to central algorithm







#### **Other Federated Tools**

#### Flimma

- Differential expression analysis based on the limma voom pipeline
- -URL: exbio.wzw.tum.de/flimma/
- Olga Zolotareva et al. (published in Genome Biology)

#### Partea

- Time-to-event studies
- -URL: exbio.wzw.tum.de/partea/
- Julian Späth et al.





#### **FeatureCloud Overview**





## FeatureCloud: Main Goals



- Convenience
  - -Implementation
  - Deployment
  - -Usage

- Trust
  - -Certification
  - -FL, DP, SMPC
  - Consent management

- Community
  - Developers
  - -Hospitals / MDs
  - -Patients







#### **FeatureCloud AI Store**







#### **Collaborative Workflows**







## **Testing & Debugging**







### **CLI and App Templates**

- Command line interface (pip install FeatureCloud)
- Python state machine template
- Example apps
- Available on GitHub (github.com/FeatureCloud)





## **App Communication**



#### App communication

- -Star-based federated learning
- -Peer-to-peer (P2P)
- -Secure multi-party computation (SMPC)





#### **System Architecture**







### **Additive Secret Sharing**







#### **Evaluation Workflow**







### **Evaluation Results (Classification)**







#### **Evaluation Results (Runtime)**







## **Future Steps**

- Integrate further privacy techniques (e.g., DP)
- Integrate compute modules (e.g., gradient descend)
- Connect to a federated database
- Improve certification process



## 1st Federated Learning and Privacy-Preserving AI Hackathon

- When: Mon, 20.06.2022 Fri, 24.06.2022
- Where: Hamburg, Germany
- Who: Undergraduate/college students, Ph.D. students and postdocs interested in the development and application of federated learning and privacy-preserving AI
- Requirements: First experience in programming (preferably Python) necessary. Knowledge of statistics/machine learning and biomedical data is helpful.
- Registration: featurecloud.ai/hackathon
- Limited number of travel scholarships covering accommodation and flight costs!



FeatureCloud



# Thank you!



26